Parent and Teacher SNAP-IV Ratings of Attention Deficit/Hyperactivity Disorder Symptoms: Psychometric Properties and Normative Ratings from a School District Sample

Regina Bussing, M.D., M.S.H.S.[1], **Melanie Fernandez, Ph.D.**[1], **Michelle Harwood, Ph.D.**[1], **Wei Hou, Ph.D.**[1], **Cynthia Wilson Garvan, Ph.D.**[1], **James M. Swanson, Ph.D.**[2], and **Sheila M. Eyberg, Ph.D.**[1]

[1]University of Florida

[2]University of California at Irvine

Abstract

To examine psychometric properties of the SNAP-IV, parent ($N = 1,613$) and teacher data ($N = 1,205$) were collected from a random sample of elementary school students in a longitudinal study on detection of attention deficit hyperactivity disorder (ADHD). Reliability, factor structure, predictive validity, and effect sizes (*ES*) for differences in ratings across age, gender, and race were examined. Performance as a screening and diagnostic tool was evaluated through calculation of likelihood ratios (*LR*) and posttest probabilities. Reliability of the parent and teacher SNAP-IV was acceptable. Factor structure was consistent with a two-factor solution of ADHD symptoms and a third ODD factor. Parent and teacher scores varied significantly by gender and poverty status ($d = .49$ to $.56$), but not by age; only teacher scores varied by race ($d = .25$ to $.55$). SNAP-IV parent and teacher ratings satisfactorily distinguished children by increasing levels of ADHD concerns, but only parent ratings by diagnostic status. Parent SNAP-IV scores above 1.2 increased the probability of concern (*LR > 10*) and above 1.8 of ADHD diagnosis (*LR > 3*). Teacher hyperactivity/impulsivity scores above 1.2 and inattention scores above 1.8 increased the probabilities of concern (*LR = 4.2* and *> 5*, respectively), but teacher SNAP-IV scores did not usefully change the probability of diagnosis. Further research should address reasons for higher teacher scores for African American children and the differences in measurement models by race.

Keywords

Behavior rating scales as assessment tools for diagnosing attention deficit hyperactivity disorder (ADHD) based on the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* have been used for many years, starting with the assessment of symptoms listed in the DSM-III (American Psychiatric Association, 1980) manual (Swanson, Sandman, Deutsch, & Baren, 1983). These scales have changed along with revisions in the DSM to match definitions of ADHD as a three-dimensional construct in DSM-III (APA, 1980), a unidimensional construct in the DSM-III-R (APA, 1987), and the current two-dimensional construct in the DSM-IV (APA, 1994) manual (Pillow, Pelham, Hoza, Molina, & Stultz, 1998). The rating scales are comparable in content (using either the exact DSM symptom

Correspondence should be directed to Regina Bussing, M.D., M.S.H.S., at Box 100234 UFHSC Gainesville, FL 32610-0234. rbussing@psychiatry.ufl.edu.

descriptions or variants slightly reworded to improve readability) and measurement approaches (four-point rating intervals) but differ in the assessment of comorbid disorders. Some focus on ADHD only, such as the ADHD Rating Scale IV (DuPaul et al., 1998; DuPaul et al., 1997) or the DSM-IV ADHD Rating Scale (Gomez, Harvey, Quick, Scharer, & Harris, 1999). Still others, such as the parent and teacher versions of the Vanderbilt ADHD Diagnostic Rating scales (Wolraich, Feurer, Hannah, Baumgaertel, & Pinnock, 1998; Wolraich et al., 2003), assess both externalizing and internalizing disorders and include items targeting impairment as well.

Among available DSM-IV-based ADHD rating scales, the Swanson, Nolan, and Pelham –IV Questionnaire (SNAP-IV) has been used in many treatment studies, including the Multimodal Treatment Study for ADHD (MTA; Correia Filho et al., 2005; Swanson et al., 2001; The MTA Cooperative Group, 1999a, 1999b) and also in genetic studies (Smalley et al., 2000; Willcutt, Pennington, Chhabildas, Friedman, & Alexander, 1999), even though published psychometric properties are lacking and normative data are sparse (Collett, Ohan, & Myers, 2003). The SNAP was originally developed to assess ADHD symptoms according to the DSM-III (American Psychiatric Association, 1980) and has been updated with subsequent revisions of the DSM (see Swanson, 1992). The long form of the SNAP-IV assesses ADHD, oppositional defiant disorder (ODD), and overlapping symptoms of all other psychiatric disorders of childhood listed in DSM-IV (see Swanson, 1992), and is available at http://www.ADHD.net. A short, 26-item SNAP-IV version, also referred to as the MTA version, assesses ADHD core symptoms of hyperactivity/impulsivity and inattention, along with symptoms of ODD. Scoring instructions for the SNAP-IV are provided on the SNAP website (http://www.ADHD.net/snap-iv-instructions.pdf). Average rating indices are constructed for inattentive, hyperactive/impulsive, combined ADHD, and ODD subscales, and scores above the 95th percentile are labeled clinically relevant. However, the website cutoff scores provided for parent and teacher ratings do not provide age and gender stratification. Furthermore, the study sample used to produce these cutpoints (Gaub & Carlson, 1997) consisted predominantly of low-income Hispanic elementary school students, limiting generalizability of findings.

In a recent review of ADHD rating scales, Collett and colleagues (2003) criticized the lack of published psychometric properties and sparse normative data (Collett et al., 2003). One study using the SNAP-IV reported good to excellent internal consistency (Stevens, Quittner, & Abikoff, 1998), although interrater agreement between parents and teachers, as typical of many behavior rating scales, has been poor (Swanson et al., 2001). Factor analyses have been reported on the SNAP website (http://www.ADHD.net) but not published. Collett and colleagues concluded that the absence of age- and gender-specific normative data especially limits the usefulness of the SNAP-IV in clinical practice.

This study had two goals. The first purpose was to examine the psychometric properties of the MTA version of the parent and the teacher SNAP-IV in an epidemiologically derived student sample from a North Central Florida school district and to explore the need for age-, gender-, and race-specific normative data. The second goal was to investigate the utility of the SNAP-IV rating scale for population screening and for diagnostic assessment of ADHD symptoms.

Method

Data were gathered as part of a longitudinal study on ADHD detection and service use (Bussing et al., 2005). The study sample was derived from school registration records that identified 12,009 elementary school students enrolled in kindergarten through 5th grade in a North Central Florida public school district during academic year 1998–1999. Of these,

3,158 students were selected for Phase 1, the risk-screening phase, using a gender-stratified random selection design in which girls were over-sampled by a margin of two to one to ensure adequate representation of girls with ADHD symptoms for Phase 2. Only one child per household was eligible for Phase 1 selection to ensure participant independence. Children were eligible for the study if they lived in a household with a telephone, were not receiving special education services for mental retardation or autism, and were from Caucasian or African American backgrounds. Children from other ethnic backgrounds were excluded because they comprised less than 5% of the total student population in the school district. Analyses were adjusted for this sample design and differential response, namely participation rates were higher for children who were Caucasian, socioeconomically advantaged, and identified as qualifying for gifted education during Phase 1, using analytic weights computed in a procedure outlined by Aday and described in detail elsewhere (Aday, 1996). Based on Phase 1 findings, children were eligible for Phase 2, the diagnostic and services assessment phase, if: (a) they were diagnosed with or undergoing treatment for ADHD; (b) either their parents or teachers had voiced concern about possible ADHD; or (c) either their parents or teachers had voiced behavioral (but not specific ADHD) concern, and they received elevated scores on the SNAP-IV parent rating scale.

Data collection procedures

Following informed consent procedures approved by the Institutional Review Board of the University of Florida and by the school district research director, telephone interviews with parents were conducted from October through December 1998. Telephone interview included inquiries into the child's health status, parental knowledge and attitudes about ADHD, a structured ADHD detection and service use assessment, and SNAP-IV ratings. Contact was made with 64% ($n = 2,035$) of the eligible sample ($N = 3,158$). The cooperation rate was 79%, yielding 1,613 completed parent interviews. During the interview, parent permission was obtained from 96% of respondents to ask the child's homeroom teacher to complete the teacher SNAP-IV rating scale. Of the 1,549 mailed teacher questionnaires, 78% ($n = 1,205$) were completed and returned. Of 382 children eligible for Phase 2, 266 (70%) participated, 83 (22%) refused, and 32 (8%) could not be contacted. Parents and children participating in Phase 2 completed diagnostic interviews, self-report measures, and services assessments. Phase 2 participants did not differ from Phase 1 participants in demographic characteristics.

Parental Concern Level

As part of the structured Phase 1 interview, parents were asked whether there had been any general concerns that their child might have an emotional or behavioral problem; whether parents or school staff suspected that their child had ADHD, attention deficit disorder, ADD, attention deficit, or hyperactivity; or whether their child had ever had a professional evaluation for ADHD. Eight percent of children ($n = 127$) had reportedly received a professional ADHD diagnosis and were labeled "diagnosed ADHD." This percentage is close to the estimated 9% reported for the State of Florida [Centers for Disease Control and Prevention (CDC), 2005]. For 191 children (12%), either the parents or school staff had voiced suspicion of ADHD, but no diagnostic assessment had been conducted. These children were labeled "suspected ADHD." The children with diagnosed and suspected ADHD ($n = 318$, 20%) together were classified as "ADHD Concern." For another 439 children (27%), parents and/or school staff had voiced concerns about the child's emotions or behavior, without suspicion or diagnosis of ADHD, and these children were classified as "General Concern." The remaining 856 children (53%) were categorized as "No Concern."

Measures

SNAP-IV—The MTA version of the SNAP-IV (Swanson et al., 2001) was used to obtain ratings from two sources, parents and teachers. The 26 items of the MTA SNAP-IV include the 18 ADHD symptoms (9 for inattentive, 9 for hyperactive/impulsive) and 8 ODD symptoms specified in the DSM-IV. Items are rated on a 4-point scale from (0) *not at all* to (3) *very much.* Average rating-per-item (ARI) subscale scores for both parent and teacher scales are calculated for the inattention, hyperactivity/impulsivity, and opposition/defiance domains, resulting in six SNAP-IV subscale scores that can range from 0 to 3, abbreviated subsequently as P-Inatt, P-Hyp/Imp, P-Odd, T-Inatt, T-Hyp/Imp and T-Odd.

Sociodemographic characteristics—Information about gender, age, race, grade level, and lunch subsidy status was obtained from school district administrative records. Based on federal government guidelines involving family income, lunch status was identified as subsidized or non-subsidized, with subsidized lunch corresponding to lower socioeconomic status (SES). SES scores also were calculated using the Hollingshead (1975) Four Factor Index, which ranges from 8 (lowest social strata) to 66 (highest strata), based on parental education and occupation (Hollingshead, 1975). Sociodemographic characteristics of this representative school district sample and of the Phase 2 sample are shown in Table 1. Teacher ratings reflect slightly higher questionnaire completion for economically advantaged children, (78% versus 72%), $\chi^2 (1, 1613) = 7.82$, $p < .01$, Caucasians (76% versus 71%), $\chi^2 (1, 1613) = 5.66$, $p < .05$, and children in the lowest grades (77% versus 72%), $\chi^2 (1, 1613) = 4.71$, $p < .05$, than their disadvantaged, African American, or higher-grade peers.

Diagnostic Interview Schedule for Children, Parent Version (NIMH DISC-IV-P; (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000)—For Phase 2 participants, diagnoses of ADHD, ODD, and conduct disorder (CD) were made with the DISC-IV-P, which uses criteria contained in the DSM-IV (American Psychiatric Association, 1994) and inquires about symptoms and impairment in both home and school settings. Computer-assisted DISC-IV-P interviews were conducted by the principal investigator, co-investigator, and three senior psychology graduate students, after intensive training sessions and establishment of interrater reliability (99%). In its earlier versions, the DISC was shown to have moderate to substantial test-retest reliability and internal consistency (Fisher et al., 1993; Jensen et al., 1995; Piacentini et al., 1993; Schwab-Stone, Fallon, Briggs, & Crowther, 1994). Despite its greater length and complexity, the test-retest reliability of the DISC IV-P compares favorably with the earlier versions (Shaffer et al., 2000). Cronbach's alpha of .93 for the DISC-IV-P ADHD module has been reported (Wolraich et al., 2003).

Data Analysis

Psychometric properties of the SNAP-IV were examined in terms of internal reliability, interrater reliability, and factor structure. These analyses were conducted using analytic weights described above. To assess internal reliability, coefficient alphas for parent and teacher ratings were calculated for the combined 26 items and for the three subdomains of the SNAP-IV (inattention, hyperactivity/impulsivity, and ODD). Interrater reliability was calculated using Pearson correlations between parent and teacher subscale scores. Confirmatory factor analysis (CFA) was performed using the method of random parceling, a method that reduces a large number of indicators to yield a better model fit (Little, Cunningham, Shahar, & Widaman, 2002). We examined 3-factor (domains of inattention, hyperactivity/impulsivity, and ODD) and 4-factor models (domains of inattention, hyperactivity, impulsivity, and ODD), for both the teacher and the parent SNAP-IV data. Three items were randomly selected from the same domain, with equal probability for each item, and the average of the three items represented a parcel to be loaded onto the latent

variable. For ODD, one parcel aggregated two items. In the 4-factor model, the domain of impulsivity consisted of only 3 items, so these items were averaged into a parcel without random selection. A Box-Cox power transformation was used to normalize the SNAP-IV score distributions (Box & Cox, 1964).

Cohen's *d* values, defined as the standardized mean difference between groups (Cohen, 1988), were calculated to explore the meaningfulness of differences among subpopulations categorized by age, gender, race/ethnicity, and poverty. These calculations were performed because our large sample size produced statistically significant independent contributions on MANOVA analyses for one-half of the age, all of the gender, all but two (P-Inatt, P-Hyp) of the race/ethnicity, and all of the poverty estimates. Cohen (1988) suggested interpreting effect sizes of 0.2 as small, 0.5 as medium, and 0.8 as large.

To assess predictive validity and the utility of SNAP-IV ratings as a screening and diagnostic tool, analyses were performed separately for the Phase 1 representative school district sample, examining whether a child had caused parental concern, and the Phase 2 high-risk sample, examining whether a child in the high risk category met DSM-IV criteria for ADHD. All analyses conducted to examine screening performance were unweighted. Phase 1 participants were grouped by concern level, obtained during the parental screening interview, into "No Concern" ($n = 856$; 53%) and "Any Concern" ($n = 757$; 47%), with the "Any Concern" group further separated into a "General Concern" group ($n = 439$; 27%) and an "ADHD Concern" group ($n = 318$; 20%). Phase 2 participants were separated into "Diagnosis Negative" and "Diagnosis Positive" groups, based on their DISC-IV-P results. Children were classified as "Diagnosis Positive" for *inattention* if they displayed at least six of the nine DSM-IV inattention symptoms, indicating that they met criteria for the combined or the predominantly inattentive subtype of ADHD. Similary, they were classified as "Diagnosis Positive" for *hyperactivity/impulsivity* if they displayed six of the nine DSM-IV hyperactivity/impulsivity symptoms. This approach was chosen because it corresponds to the two SNAP-IV inattention and hyperactivity factors, whereas using the DSM-IV subtypes would have yielded three diagnostic categories, the combined, predominantly inattentive, and predominantly hyperactive subtypes.

Predictive validity of the parent and teacher SNAP-IV ADHD subscales were examined by comparing children's scores across the three Concern Level groups identified in Phase 1, and across the two Diagnostic Status groups (ADHD positive and negative, based upon DISC-IV-P results) identified in Phase 2, using the GLM procedure and adjusting for multiple comparisons with the Tukey-Kramer method.

The area under the Receiver Operating Characteristics (ROC) curve was calculated to measure the "accuracy" of SNAP-IV predictions. Benchmarks for interpreting accuracy have been suggested by Swets (Swets, 1988). ROC values between .5 and .7 represent "low" accuracy, meaning the true-positive proportion is not much greater than the false-positive proportion anywhere along the curve. ROC values between about .7 and .9 are classified as "useful," and values above .9 are classified as "high" accuracy (Swets, 1988).

We also conducted hierarchical logistic regression analyses to evaluate the incremental validity of teacher SNAP-IV scores, above and beyond that provided by parent SNAP-IV scores alone, in predicting Concern status (Phase 1 participants) or ADHD diagnosis on the DISC-IV-P (Phase 2 participants) in this sample. For these models, parent inattention and hyperactivity/impulsivity scores were entered separately in Step 1, and the corresponding teacher scores were entered in Step 2.

Likelihood Ratios (*LR*) (*LR* = sensitivity/1 – specificity) were calculated for three SNAP-IV score ranges with respect to "actual values" of Concern status and DISC-IV-P ADHD

diagnosis. *LR*s > 0.5 and < 2 are not considered useful; *LR*s from 0.2 – 0.5 or 2 – 5 are considered useful; and *LR*s < 0.2 or > 5 tend to produce large changes in post-test probability. Using Bayes Theorem, posttest probabilities were calculated for three base rates, intended to correspond to community (5%), pediatric practice (20%) and mental health specialty settings (50%). To obtain posttest odds (or probability that a person meets the gold standard given a positive test result), the *LR* was multiplied by the pretest odds. Concern and ADHD Diagnosis were separately considered the "gold standards" for these calculations.

Results

Reliability

Coefficient alpha for overall parent ratings was .94. For the inattentive, hyperactive/ impulsive, and ODD subdomains, coefficient alphas were .90, .79, and .89, respectively. Coefficient alpha for overall teacher ratings was .97, with subdomain alphas of .96, .92, and .96, respectively. There were no significant variations in internal consistency by gender or race for either parent or teacher SNAP-IV ratings. Interrater reliability between parent and teacher ratings was .49 for inattention, .43 for hyperactivity/impulsivity, and .47 for ODD, and all were statistically significant ($p < .001$). Internal consistency estimates did not increase with removal of any item; specifically, ranges of alpha (if item deleted) for the inattentive, hyperactive/impulsive, and ODD subdomains were .88 to .89, .76 to .80, and .87 to .90 for parents and .95 to .96, .91 to .92, and .95 to .96 for teachers, respectively.

Factor Analysis

Confirmatory Factor Analysis (CFA) results indicated a better fit for the 3-factor model than the 4-factor model for parent and for teacher data. We used the ECVI (Expected Cross-Validation Index) and AIC (Akaike's Information Criterion) to compare models, which are both useful indices for comparing non-nested models. The ECVI values for the 3-factor and 4-factor models were .091 and .120, respectively, for parent models, and .156 and .173 for the teacher models. Corresponding AIC values were 56.632 and 103.133, and 96.023 and 116.695. Thus, we selected the 3-factor model for further presentation and to test the fit of a series of 2-group CFA models with subgroups based on gender and race. CFA results for the 3-factor models are depicted in the Figure; teacher results are shown in parentheses. CFA models resulted in satisfactory fit indices. For the SNAP-IV parent model, Goodness of Fit = .99, Root Mean Square = .05, Comparative Fit Index = .94, and Non-normed Index = .91; the corresponding teacher estimates were .97, .06, .91 and .86. Tests of the equality of factor loadings across groups (Caucasian/African American and males/females) in teacher scores indicated that factor loadings were not equivalent (race X^2 (6) = 20.63, gender X^2 (6) = 27.78, *p*-values < .05). However, the range of differences in standardized factor loadings tended to be trivial (race .01 – .05; gender .00 – .02). The factor loadings were equivalent across groups (Caucasian/African American and males/females) in parent scores (race X^2 (6) = 4.87, gender X^2 (6) = 13.39, p-values > .05).

Mean Parent and Teacher SNAP-IV Scores

Tables 2 and 3 depict SNAP-IV subscale scores stratified by gender and age, as well as gender, race, and age. Almost all subscale scores were below 1.0 (equivalent to a rating below "just a little") and reflected the expected ratings of symptoms of ADHD (i.e., signs of psychopathology) in a population sample in which by definition the large majority of individuals should not manifest the symptoms.

Exploration of Age, Gender, Race, and Poverty Effects

Cohen's *d* estimates for age differences wereseparated into three age groups, 5- to 7-year-olds, 8- to 10-year-olds, and 11-year-olds, to make results comparable to DuPaul et al., 1997. These estimates showed a small effect size for parent inattention ratings comparing 8- to 10-year-olds (.33) and 5- to 7-year-olds (.43) with 11-year-olds, with higher ratings for the oldest group. All other age group comparisons showed effect sizes below 0.2. Cohen's *d* for gender differences was small for parent (P-Inatt = .40; P-Hyp = .37; P-Odd = .31) and small to medium-sized for teacher (T-Inatt = .49; T-Hyp = .40; T-Odd = .29) ratings. On all six parent and teacher comparisons, boys received significantly higher SNAP-IV ratings than girls ($p < .0001$). Cohen's *d* estimates for SNAP-IV score differences by race were negligible for parents (P-Inatt = .11; P-Hyp = .17; P-Odd = .06) and in the medium-sized range for teachers (T-Inatt = .56; T-Hyp = .49; T-Odd = .51), consistently showing higher teacher ratings for African American than Caucasian children ($p < .0001$). Cohen's *d* estimates for SNAP-IV score differences by poverty status were small for parents (P-Inatt = .25; P-Hyp = .33; P-Odd = .28) and in the medium-sized range for teachers (T-Inatt = .55; T-Hyp = .47; T-Odd = .48), consistently showing higher teacher ratings for children in poverty than their higher SES peers ($p < .0001$). Because none of the effect size estimates were of large size, subsequent analyses were not stratified by age, gender, race, or poverty.

Predictive Validity: SNAP-IV Subscale Scores by Concern Level and DSM-IV Diagnostic Status

As shown in Table 4, average parents and teacher SNAP-IV subscale scores increased significantly with rising ADHD concern, and *p* values of Tukey Cramer adjusted pair-wise comparisons between Concern Levels were all < 0.0001. These findings support the SNAP-IV's predictive validity for Concern Levels in a community sample. Furthermore, both SNAP-IV parent subscale scores were significantly higher for children who met ADHD criteria on the DISC-IV than for those who did not ($p < .0001$). Teacher SNAP-IV inattention scores differed at trend level by diagnostic status (1.20 versus 1.46, $p = .052$), whereas hyperactivity/impulsivity scores did not differ.

Incremental Predictive Validity of SNAP-IV Teacher Scores

Results of the hierarchical logistic regression analyses indicated that adding teacher inattention and hyperactivity/impulsivity subscale scores significantly improved the prediction of Concern Level (chi square = 39.32, $df = 2$, $p < .0001$) compared to using parent SNAP-IV scores alone. However, teacher SNAP-IV ratings did not add to the prediction of diagnostic status for either inattentive or hyperactive presentations. The poorer diagnostic predictive validity of teacher SNAP-IV scores may reflect shared informant variance because parent-based DISC-IV-P interviews were used to establish diagnoses.

SNAP-IV Score Performance as a Screening and Diagnostic Tool

Parent and teacher SNAP-IV subscale ratings were found to have useful accuracy (P-Inatt = .85; P-Hyp = .80; T-Inatt = .74; T-Hyp = .71) for distinguishing children eliciting behavioral/emotional concerns from those who did not. In contrast, only parent SNAP-IV subscale ratings were found to have useful accuracy (P-Inatt = .72; P-Hyp = .76; T-Inatt = .61; T-Hyp = .58) for differentiating high-risk children who met DSM-IV criteria for ADHD in Phase 2 from those who did not.

Table 5 illustrates how various score ranges for SNAP-IV subscales performed in "ruling in" or "ruling out" Concern Level and Diagnostic status. Parent subscale scores of inattention and of hyperactivity/impulsivity above 1.2 produced large increases in the likelihood (*LRs = 21.4 and 13.1, respectively)* that a child prompted Concern. Scores above 1.2 on both

teacher SNAP-IV subscales produced useful increases in the likelihood of Concern (*LR*s > 2.0 and < 5.0), and teacher inattention scores but not hyperactivity/impulsivity scores above 1.8 raised the *LR* from useful to large (*LR* > 5). Parent SNAP-IV inattention scores above 1.8 and hyperactivity/impulsivity scores above 2.4 produced useful increases in the likelihood that a child would meet DISC-IV-P diagnostic criteria for inattention or for hyperactivity/impulsivity, respectively. Increasing teacher SNAP-IV subscale scores were not associated with increased likelihood of an ADHD diagnosis.

The last three columns of Table 4 also demonstrate for three hypothetical base rate settings the corresponding posttest probabilities that a child would elicit Concern or meet ADHD diagnostic criteria, given various parent and teacher SNAP-IV subscale score ranges.

Discussion

This study presents initial psychometric and school-district-based standardization data for the parent and teacher SNAP-IV rating scale. The findings suggest no need for age-specific normative cutpoints among elementary school students. In addition, gender and race differences in SNAP-IV scores were not large and, therefore, support the unstratified cutpoint approach offered on the SNAP website (www.ADHD.net). We caution that our sample, although large and representative of the school district, was restricted to one school district in North Florida with high poverty rates and limited diversity, and findings may not generalize to other regions. Geographic variations in ADHD prevalence may require separate norms for distinct regions of the U.S. (Stevens, Harman, & Kelleher, 2004).

Internal consistency, item selection, and factor structure of the SNAP-IV were found acceptable and consistent with the constructs put forth in the DSM-IV, which would be expected given the similarity between the SNAP-IV items and the DSM-IV criterion symptoms.

Careful analysis of SNAP scores by child age did not support the notion of developmental amelioration of ADHD-related behaviors during the elementary school years. This is consistent with other studies reporting negligible to small age effects. For example, differences in the Conners' DSM inattention and hyperactivity/impulsivity scale scores between ages 5 and 11 range from effects sizes of 0.06 to 0.23 for parents and from 0.07 to 0.16 for teachers (Conners, 1997). However, our cross-sectional design precludes firm conclusions about age effects.

Parent and teacher ratings of hyperactivity/impulsivity, inattention, and ODD were higher for boys than girls, but the differences were small to medium-sized. Our results for the SNAP-IV are commensurate with previous findings of gender discrepancies in perceived severity of ADHD symptom expression according to parent and teacher ratings of DSM ADHD behavior (DuPaul et al., 1998; DuPaul et al., 1997; Reid et al., 2000). Although gender-specific norms may attenuate the influence of halo effects, they also may artificially alter prevalence rates in ADHD (Silverthorn, Frick, & Kuper, 1996), and they imply that girls have to meet less stringent standards for diagnosis than boys (Robinson, Eyberg, & Ross, 1980). Current DSM-IV terminology for ADHD has been viewed as more typical of boys than girls, and some evidence suggests that girls manifest ADHD symptoms differently than boys (e.g., girls more frequently changing friends impulsively or without thinking) (Ohan & Johnston, 2005). Our study findings do not support the use of gender-specific norms among elementary school age students on the SNAP-IV and factor analysis results only demonstrated trivial differences in factor loadings that did not appear meaningful. These findings suggest relative equivalence rather than differences in measurement models by gender.

Our findings suggest that parents and teachers differ in their reports of child behavior. Differences in parent SNAP-IV ratings of African American and Caucasian children produced a negligible to small effect size, whereas teacher ratings produced systematically higher scores for African American than Caucasian children, of medium effect size. Similar results for parent ratings of ADHD have been found in other studies (DuPaul et al., 1998). Race and ethnic differences in teacher ratings of ADHD symptoms have been reported in other studies as well (Epstein, March, Conners, & Jackson, 1998; Floyd, Rayfield, Eyberg, & Riley, 2004; DuPaul et al., 1997; Epstein et al., 2005; Robert Reid, Casat, & Norton, 2001; Robert Reid et al., 2000). In response to findings that teachers but not parents reported race differences in ADHD symptoms our differential teacher response rate by race must be considered. We had slightly lower questionnaire completion for African American (71%) than Caucasian (76%) children. If teachers were less likely to complete SNAP ratings for asymptomatic African American children, the overall group symptom level relative to Caucasian children may have been inflated artificially.

It is also important to consider potential rater effects (Robert Reid et al., 2000). In the MTA study (The MTA Cooperative Group, 1999a), observed ADHD symptoms were significantly correlated with teacher ratings and were higher for African American than Caucasian children, although the ethnic differences were reduced when statistically controlling for the behavior of an average child in the classroom (Epstein et al., 2005).

Use of unstratified cutpoints has notable implications, particularly for the greater number of African American relative to Caucasian children who will screen positive for inattention and hyperactivity/impulsivity. However, it remains impossible to conclude whether differences in teacher ratings of medium effect size were based on rater or halo effects, contextual variables in the classroom, or true differences in child behavior. Thus, we maintain the appropriateness of using unstratified norms with an understanding that teacher SNAP-IV ratings must be interpreted with caution and are not diagnostic in and of themselves. Notably, given the difference in factor loadings for models tested by race, additional steps toward achieving measurement equivalence of the ADHD construct should be taken.

Study findings suggest that the SNAP-IV adequately discriminates children with varying levels of ADHD concern, but more caution is warranted when using it as a diagnostic tool. As a screening measure for emotional/behavioral concerns, the SNAP-IV performs adequately, with modest parent or teacher SNAP-IV subscale scores elevations predicting useful increases in the likelihood of Concern. Thus, clinicians in moderate to high prevalence settings can be fairly certain that parent SNAP-IV scores above 1.2 and teacher scores above 1.8 signal behavioral concerns that the child merits a diagnostic assessment for ADHD. For purposes of differentiating ADHD positive from negative cases, parent hyperactivity/impulsivity scores above 1.8 and parent inattention scores above 2.4 usefully (LR 3 – 5) increase the probability of ADHD diagnosis, but with lower post-test probabilities than achieved for identifying Concern. This finding is of particular importance for primary care settings, because ADHD prevalence rates are lower than in specialty mental health settings (Brown et al., 2001; Hoagwood, Kelleher, Feil, & Comer, 2000) and many pediatricians and family practitioners rely on behavior rating scales during their diagnostic assessments. For example, in a recent national primary care survey Chan and colleagues found that 58% of practitioners reported using formal criteria to diagnose ADHD, but only 28% indicated using DSM-IV criteria while close to 70% relied on rating scales (Chan, Hopkins, Perrin, Herrerias, & Homer, 2005). Study findings clearly caution against reliance on the SNAP-IV as a diagnostic instrument, yet, this caveat also applies to other ADHD behavioral rating scales. Based on carefully designed research examining the predictive validity of parent and teacher ratings of ADHD among a clinic-referred sample, Power et al. (1998) cautioned that even using optimal combined parent and teacher rating cutoffs would

results in about one-quarter of children with combined subtype and over 40% of those with inattentive subtype being missed.

We acknowledge limitations of our data for interpreting race differences in the current study. We did not obtain data on individual teachers' race or ethnicity (the school district reports that 68% of teachers are Caucasian, 28% African American, and 3% Hispanic) nor on contextual factors related to schools or classrooms, which might explain some possible sources of bias. We recognize that elevated scores for African-Americans on the teacher SNAP-IV may be due to the association of race status and high poverty rates, with poverty identified as a risk factor for ADHD (Biederman et al., 1995). Future studies should incorporate teacher ratings, classroom behavioral observations, and teacher and observer ethnicity to help distinguish between true race or ethnic differences in ADHD symptoms and rater effects. Several other limitations should be noted when considering study findings. In addition to the regional nature of the study sample, design limitations of our two-stage screening approach need to be considered. Although Phase 1 data represent the school district students, Phase 2 data reflect the high-risk criteria (parental concerns, suspicion of ADHD, and past ADHD diagnosis or treatment) selected for purposes of studying ADHD detection processes in this population. Thus, screener performance in the high-risk sample may change with altered risk definitions and reported Likelihood Ratios, based on this specific sample, may not generalize to other settings. Of note, the choice of Concern as a "gold standard" to assess screening performance is less established than using measures with known validity; however, studies have demonstrated that carefully elicited parental concerns can reliably identify children with mental health problems (Glascoe, 2003). In addition, interpretations of the teacher SNAP-IV need to consider that the DISC-IV-P relies on parental reports of school behavior and functioning, therefore teacher perceptions are underrepresented in the determination of DSM-IV diagnoses.

While our findings need to be interpreted with these limitation in mind, in summary, this study suggests that the SNAP-IV questionnaire has acceptable internal consistency and item selection, and a factor structure consistent with the two-factor solution of ADHD symptoms, and a third ODD factor. Results imply that the effect sizes of differences in parent and teacher subscale scores by gender and poverty are small to medium-sized and that only teacher ratings show medium-sized effects by race. Further research needs to address reasons for higher teacher scores in African American children, as well as the differences in measurement models by race suggested by our factor analysis results. While study results suggest that SNAP-IV parent and teacher ratings satisfactorily distinguish children with different levels of ADHD concerns, findings caution against using the SNAP-IV for diagnostic purposes.

Acknowledgments

References

Aday, LA. Designing and conducting health surveys. San Francisco: Jossey-Bass Publishers; 1996.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed.. Washington, D.C.: Author; 1980.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed., revised. Washington, D.C.: Author; 1987.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed.. Washington, D.C.: Author; 1994.

Baren M, Swanson JM. How not to diagnose ADHD. Contemporary Pediatrics. 1996; 13:53–64.

Biederman J, Milberger S, Faraone SV, Kiely K, Guite J, Mick E, et al. Family-environment risk factors for attention-deficit hyperactivity disorder. A test of Rutter's indicators of adversity. Archives of General Psychiatry. 1995; 52:464–470. [PubMed: 7771916]

Box GEP, Cox DR. An analysis of transformations. Journal of the Royal Statistical Society. 1964; 26:211–252.

Bussing R, Zima BT, Mason D, Hou W, Garvan CW, Forness S. Use and persistence of pharmacotherapy for elementary school students with attention-deficit/hyperactivity disorder. Journal of Child and Adolescent Psychopharmacology. 2005; 15:78–87. [PubMed: 15741789]

Centers for Disease Control and Prevention (CDC). Mental health in the United States. Prevalence of diagnosis and medication treatment for attention-deficit/hyperactivity disorder--United States, 2003. MMWR Morbidity Mortality Weekly Reports. 2005; 54:842–847.

Chan E, Hopkins MR, Perrin JM, Herrerias C, Homer CJ. Diagnostic practices for attention deficit hyperactivity disorder: a national survey of primary care physicians. Ambulatory Pediatrics. 2005; 5:201–208. [PubMed: 16026184]

Cohen, J. Statistical power analysis for the behavioral sciences. Hillsdale, New York: Lawrence Erlbaum Associates; 1988.

Collett BR, Ohan JL, Myers KM. Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry. 2003; 42:1015–1037. [PubMed: 12960702]

Conners, KC. Connors' Rating Scales-Revised Technical Manual. North Tonawanda, NY: Multi-Health Systems, Inc.; 1997.

Correia Filho AG, Bodanese R, Silva TL, Alvares JP, Aman M, Rohde LA. Comparison of risperidone and methylphenidate for reducing ADHD symptoms in children and adolescents with moderate mental retardation. Journal of the American Academy of Child and Adolescent Psychiatry. 2005; 44:748–755. [PubMed: 16034276]

DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, Ikeda MJ, McGoey KE. Parent ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. Journal of Psychopathology and Behavioral Assessment. 1998; 20:83–102.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ. Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. Psychological Assessment. 1997; 9:436–444.

Epstein JN, March JS, Conners CK, Jackson DL. Racial differences on the Conners Teacher Rating Scale. Journal of Abnormal Child Psychology. 1998; 26:109–118. [PubMed: 9634133]

Epstein JN, Willoughby M, Valencia EY, Tonev ST, Abikoff HB, Arnold LE, et al. The role of children's ethnicity in the relationship between teacher ratings of attention-deficit/hyperactivity disorder and observed classroom behavior. Journal of Consulting and Clinical Psychology. 2005; 73:424–434. [PubMed: 15982140]

Fisher PW, Shaffer D, Piacentini J, Lapkin J, Kafantaris V, Leonard H, et al. Sensitivity of the Diagnostic Interview Schedule for Children, 2nd edition (DISC-2.1) for specific diagnoses of children and adolescents. Journal of the American Academy of Child and Adolescent Psychiatry. 1993; 32:666–673. [PubMed: 8496131]

Floyd EM, Rayfield A, Eyberg SM, Riley JL. Psychometric properties of the Sutter-Eyberg Student Behavior Inventory with rural middle school and high school children. Assessment. 2004; 11:1–9.

Gaub M, Carlson CL. Behavioral characteristics of DSM-IV ADHD subtypes in a school-based population. Journal of Abnormal Child Psychology. 1997; 25:103–111. [PubMed: 9109027]

Glascoe FP. Parents' evaluation of developmental status: how well do parents' concerns identify children with behavioral and emotional problems? Clinical Pediatrics. 2003; 42:133–138. [PubMed: 12659386]

Gomez R, Harvey J, Quick C, Scharer I, Harris G. DSM-IV AD/HD: confirmatory factor models, prevalence, and gender and age differences based on parent and teacher ratings of Australian

primary school children. Journal of Child Psychology and Psychiatry. 1999; 40:265–274. [PubMed: 10188709]

Hollingshead. Four factor index of social class. Yale University: Department of Sociology; 1975.

Jensen P, Roper M, Fisher P, Piacentini J, Canino G, Richters J, et al. Test-retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1). Parent, child, and combined algorithms. Archives of General Psychiatry. 1995; 52:61–71. [PubMed: 7811163]

Little TD, Cunningham WA, Shahar G, Widaman KF. To parcel or not to parcel: Exploring the question, weighing the merits. Structural Equation Modeling. 2002; 9:151–173.

Ohan JL, Johnston C. Gender appropriateness of symptom criteria for attention-deficit/hyperactivity disorder, oppositional-defiant disorder, and conduct disorder. Child Psychiatry and Human Development. 2005; 35:359–381. [PubMed: 15886870]

Piacentini J, Shaffer D, Fisher PW, Schwab-Stone M, Davies M, Gioia P. The Diagnostic Interview Schedule for Children--Revised Version (DISC--R): III. Concurrent criterion validity. Journal of the American Academy of Child and Adolescent Psychiatry. 1993; 32:658–665. [PubMed: 8496130]

Pillow DR, Pelham WE Jr, Hoza B, Molina BS, Stultz CH. Confirmatory factor analyses examining attention deficit hyperactivity disorder symptoms and other childhood disruptive behaviors. Journal of Abnormal Child Psychology. 1998; 26:293–309. [PubMed: 9700521]

Power TJ, Doherty BJ, Panichelli-Mindel SM, Karustis JL, Eiraldi RB, Anastopoulos AD, et al. The predictive validity of parent and teacher reports of ADHD symptoms. Journal of Psychopathology and Behavioral Assessment. 1998; 20:57–81.

Reid R, Casat CD, Norton HJ. Using behavior rating scales for ADHD across ethnic groups: The IOWA Conners. Journal of Emotional and Behavioral Disorders. 2001; 9:210–218.

Reid R, DuPaul GJ, Power TJ, Anastopoulos AD, Rogers-Adkinson D, Noll MB, et al. Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. Journal of Abnormal Child Psychology. 1998; 26:187–198. [PubMed: 9650625]

Reid R, Riccio CA, Kessler RC, DuPaul GJ, Power TJ, Rogers-Adkinson D, et al. Gender and ethnic differences in ADHD as assessed by behavior ratings. Journal of Emotional and Behavioral Disorders. 2000; 8:38–48.

Schwab-Stone M, Fallon T, Briggs M, Crowther B. Reliability of diagnostic reporting for children aged 6–11 years: A test-retest study of the Diagnostic Interview Schedule for Children--Revised. American Journal of Psychiatry. 1994; 151:1048–1054. [PubMed: 8010363]

Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. Journal of the American Academy of Child and Adolescent Psychiatry. 2000; 39:28–38. [PubMed: 10638065]

Silverthorn P, Frick PJ, Kuper K. Attention deficit hyperactivity disorder and sex: A test of two etiological models to explain the male predominance. Journal of the American Academy of Child and Adolescent Psychiatry. 1996; 25:52–59.

Smalley SL, McGough JJ, Del'Homme M, NewDelman J, Gordon E, Kim T, et al. Familial clustering of symptoms and disruptive behaviors in multiplex families with attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry. 2000; 39:1135–1143. [PubMed: 10986810]

Stevens J, Harman JS, Kelleher KJ. Ethnic and regional differences in primary care visits for attention-deficit hyperactivity disorder. Journal of Developmental and Behavioral Pediatrics. 2004; 25:318–325. [PubMed: 15502548]

Stevens J, Quittner AL, Abikoff H. Factors influencing elementary school teachers' ratings of ADHD and ODD behaviors. Journal of Clinical Child Psychology. 1998; 27:406–414. [PubMed: 9866077]

Swanson JM, Kraemer HC, Hinshaw SP, Arnold LE, Conners CK, Abikoff HB, et al. Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. Journal of the American Academy of Child and Adolescent Psychiatry. 2001; 40:168–179. [PubMed: 11211365]

Swanson JM, Sandman CA, Deutsch C, Baren M. Methylphenidate hydrochloride given with or before breakfast: I. Behavioral, cognitive, and electrophysiological effects. Pediatrics. 1983; 72:49–55. [PubMed: 6866591]

Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988; 240:1285–1293. [PubMed: 3287615]

The MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. The MTA Cooperative Group. Multimodal Treatment Study of children with ADHD. Archives of General Psychiatry. 1999a; 56:1073–1086. [PubMed: 10591283]

The MTA Cooperative Group. Moderators and mediators of treatment response for children with attention-deficit/hyperactivity disorder: the Multimodal Treatment Study of children with Attention-deficit/hyperactivity disorder. Archives of General Psychiatry. 1999b; 56:1088–1096. [PubMed: 10591284]

Willcutt EG, Pennington BF, Chhabildas NA, Friedman MC, Alexander J. Psychiatric comorbidity associated with DSM-IV ADHD in a nonreferred sample of twins. Journal of the American Academy of Child and Adolescent Psychiatry. 1999; 38:1355–1362. [PubMed: 10560221]

Wolraich ML, Feurer ID, Hannah JN, Baumgaertel A, Pinnock TY. Obtaining systematic teacher reports of disruptive behavior disorders utilizing DSM-IV. Journal of Abnormal Child Psychology. 1998; 26:141–152. [PubMed: 9634136]

Wolraich ML, Lambert W, Doffing MA, Bickman L, Simmons T, Worley K. Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. Journal of Pediatric Psychology. 2003; 28:559–567. [PubMed: 14602846]
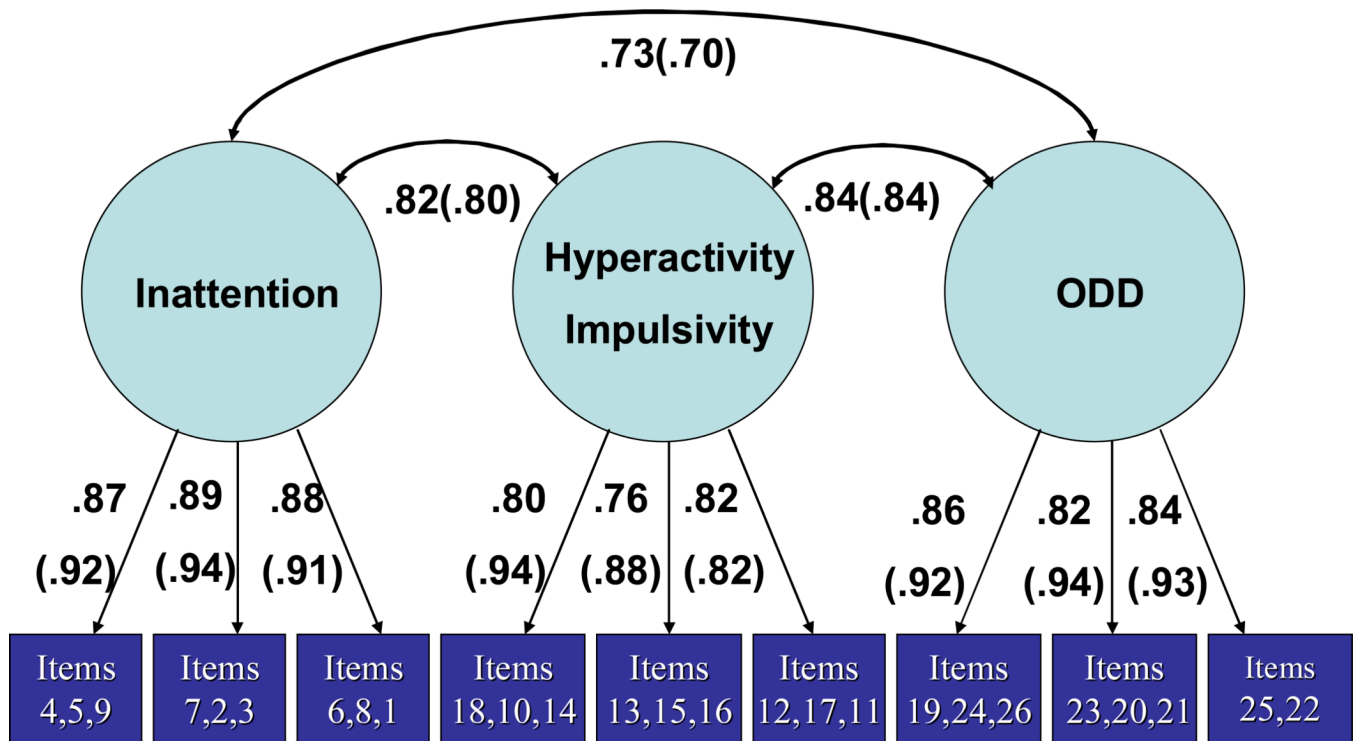
Figure.
Three-factor confirmatory model for SNAP-IV parent and (in parentheses) teacher ratings.

Table 1

Characteristics of Representative School District Samples (Parent and Teacher Rated) and a High-Risk Subgroup

	Phase 1 Parent Rated Sample[a] n (%)	Phase 1 Teacher Rated Sample[b] n (%)	Phase 2 High-Risk Sample[c] n (%)
Gender			
Male	544 (34%)	406 (34%)	136 (51%)
Female	1069 (66%)	799 (66%)	130 (49%)
Race			
Caucasian	1108 (69%)	847 (70%)	178 (67%)
African American	505 (31%)	358 (30%)	88 (33%)
Lunch status			
Subsidized	793 (49%)	568 (47%)	153 (58%)
Unsubsidized	820 (51%)	637 (53%)	113 (42%)
Grade level			
Grade K-2	830 (51%)	639 (53%)	118 (44%)
Grade 3–5	783 (49%)	566 (47%)	148 (56%)
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Age (years)	8.40 (1.59)	7.67 (1.77)	8.03 (1.73)
SES	38.12 (13.85)	38.96 (12.79)	36.00 (13.54)

Note. Females were oversampled by factor of 2:1 to assure sufficient representation in the high-risk sample.

[a] $n = 1613$;

[b] $n = 1205$;

[c] $n = 266$

Table 2

Parent SNAP-IV Scores by Years of Age in a Representative School District Sample of 1,613 Children

		Child age						
		5 (*n* = 210)	6 (*n* = 257)	7 (*n* = 294)	8 (*n* = 267)	9 (*n* = 278)	10 (*n* = 246)	11 (*n* = 61)
Child race	**Subscale**	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Boys								
African American (*n* = 167)	Inattention	.66 (.79)	.75 (.75)	1.09 (.95)	1.07 (.74)	1.12 (.90)	1.08 (1.03)	1.47 (.84)
	Hyperactive/Impulsive	.93 (.82)	.89 (.80)	1.08 (.79)	1.09 (.77)	1.02 (.94)	.81 (.79)	1.31 (.56)
	Oppositional defiant	.66 (.85)	.47 (.63)	.84 (.82)	.75 (.78)	.84 (.95)	.62 (.79)	.78 (.49)
Caucasian (*n* = 377)	Inattention	.46 (.57)	.66 (.74)	.73 (.82)	.84 (.77)	1.01 (.86)	.78 (.75)	.72 (.79)
	Hyperactive/Impulsive	.58 (.72)	.78 (.72)	.83 (.81)	.72 (.72)	.80 (.77)	.68 (.62)	.77 (.84)
	Oppositional defiant	.37 (.45)	.63 (.74)	.66 (.75)	.62 (.64)	.66 (.72)	.71 (.76)	.60 (.88)
All boys (*n* = 544)	Inattention	.52 (.64)	.69 (.74)	.84 (.87)	.92 (.77)	1.05 (.87)	.86 (.84)	.86 (.83)
	Hyperactive/Impulsive	.69 (.76)	.82 (.75)	.90 (.81)	.84 (.75)	.87 (.83)	.71 (.67)	.87 (.81)
	Oppositional defiant	.46 (.61)	.58 (.71)	.72 (.77)	.66 (.69)	.72 (.79)	.68 (.77)	.64 (.81)
Girls								
		Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
African American (*n* = 338)	Inattention	.37 (.48)	.54 (.66)	.56 (.69)	.41 (.56)	.65 (.80)	.56 (.76)	.98 (.59)
	Hyperactive/Impulsive	.66 (.65)	.58 (.56)	.67 (.66)	.40 (.40)	.73 (.71)	.41 (.52)	.99 (.73)
	Oppositional defiant	.40 (.62)	.37 (.66)	.42 (.59)	.31 (.56)	.58 (.75)	.39 (.69)	1.30 (1.06)
Caucasian (*n* = 731)	Inattention	.33 (.47)	.55 (.70)	.50 (.63)	.57 (.67)	.50 (.64)	.52 (.69)	.93 (.97)
	Hyperactive/Impulsive	.57 (.52)	.72 (.75)	.52 (.53)	.54 (.57)	.42 (.46)	.44 (.50)	.68 (.76)
	Oppositional defiant	.43 (.52)	.56 (.66)	.40 (.50)	.51 (.69)	.31 (.46)	.31 (.39)	.57 (.75)
All girls (*n* = 1069)	Inattention	.34 (.48)	.55 (.69)	.52 (.65)	.52 (.64)	.54 (.69)	.53 (.71)	.95 (.86)
	Hyperactive/Impulsive	.60 (.58)	.67 (.70)	.57 (.58)	.50 (.52)	.51 (.56)	.43 (.51)	.78 (.76)
	Oppositional defiant	.42 (.56)	.50 (.66)	.40 (.53)	.45 (.65)	.39 (.57)	.33 (.48)	.81 (.92)

Table 3

Teacher SNAP-IV Scores by Years of Age in a Representative School District Sample of 1,205 Children

		Child age						
		5 (*n* = 164)	6 (*n* = 202)	7 (*n* = 222)	8 (*n* = 183)	9 (*n* = 194)	10 (*n* = 190)	11 (*n* = 50)
Child race	**Subscale**	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Boys								
African American (*n* = 115)	Inattention	1.19 (1.03)	1.31 (.89)	1.68 (1.00)	1.31 (1.17)	1.26 (1.10)	1.78 (.85)	2.04 (1.39)
	Hyperactive/Impulsive	.97 (1.06)	1.18 (.93)	1.21 (.95)	.99 (1.20)	.87 (.97)	.91 (.78)	1.00 (.80)
	Oppositional defiant	.57 (.91)	.48 (.54)	1.04 (.97)	1.04 (1.10)	.85 (.83)	.68 (.71)	1.25 (.99)
Caucasian (*n* = 291)	Inattention	.77 (.90)	.81 (.91)	.74 (.80)	.98 (.88)	.96 (.93)	.48 (.79)	.78 (.78)
	Hyperactive/Impulsive	.47 (.71)	.57 (.67)	.42 (.63)	.59 (.69)	.62 (.78)	.25 (.56)	.61 (.72)
	Oppositional defiant	.31 (.63)	.40 (.71)	.27 (.55)	.32 (.68)	.38 (.64)	.21 (.54)	.47 (.70)
All boys (*n* = 406)	Inattention	.87 (.94)	.96 (.93)	1.02 (.96)	1.06 (.96)	1.06 (.99)	.84 (1.00)	.99 (.98)
	Hyperactive/Impulsive	.60 (.83)	.76 (.81)	.66 (.82)	.68 (.85)	.70 (.84)	.43 (.69)	.68 (.72)
	Oppositional defiant	.57 (.91)	.48 (.54)	1.04 (.97)	1.04 (1.10)	.85 (.83)	.68 (.71)	1.25 (.99)
Girls		Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
African American (*n* = 243)	Inattention	.58 (.76)	1.04 (.90)	.97 (.93)	.83 (.83)	.77 (.73)	.82 (.91)	.89 (.43)
	Hyperactive/Impulsive	.51 (.67)	.70 (.82)	.75 (.82)	.50 (.68)	.51 (.64)	.39 (.67)	.43 (.42)
	Oppositional defiant	.27 (.48)	.53 (.85)	.53 (.72)	.55 (.76)	.48 (.73)	.44 (.74)	.76 (.76)
Caucasian (*n* = 556)	Inattention	.34 (.54)	.33 (.46)	.38 (.57)	.60 (.83)	.39 (.56)	.36 (.57)	.69 (.79)
	Hyperactive/Impulsive	.31 (.52)	.27 (.51)	.22 (.48)	.36 (.64)	.21 (.45)	.20 (.41)	.35 (.51)
	Oppositional defiant	.16 (.44)	.20 (.52)	.14 (.37)	.23 (.50)	.12 (.34)	.12 (.39)	.30 (.70)
All girls (*n* = 799)	Inattention	.43 (.64)	.55 (.71)	.57 (.76)	.67 (.83)	.50 (.63)	.47 (.69)	.74 (.71)
	Hyperactive/Impulsive	.38 (.59)	.41 (.65)	.38 (.66)	.40 (.65)	.30 (.53)	.24 (.49)	.37 (.48)
	Oppositional defiant	.20 (.46)	.31 (.66)	.26 (.53)	.33 (.61)	.23 (.51)	.20 (.51)	.43 (.74)

Table 4

SNAP-IV Subscale Scores by Concern Level and Diagnosis

	ADHD Concern Level			Diagnosis	
	None	**General**	**Specific**	**Negative**	**Positive**
	Inattention Subscores				
Parent					
n	855	439	318	127	137
SD	0.30	0.64	0.88	0.75	0.77
Mean	0.24	0.74	1.57	1.15	1.84
Overall *p*-value[a]	<0.0001			<0.0001	
Teacher					
n	644	332	229	90	100
SD	0.63	0.91	1.03	1.10	0.87
Mean	0.42	1.03	1.41	1.20	1.46
Overall *p*-value	<0.0001			0.0519	
	Hyperactivity/Impulsivity Subscores				
Parent					
n	855	439	318	153	111
SD	0.32	0.62	0.87	0.72	0.80
Mean	0.32	0.76	1.46	1.08	1.83
Overall *p*-value	<0.0001			<0.0001	
Teacher					
n	640	330	227	113	76
SD	0.55	0.75	0.95	0.90	0.88
Mean	0.28	0.64	0.98	0.90	1.01
Overall *p*-value	<0.0001			0.4068	

Notes. Weighted for sample design and non-participation.

[a]P values using GLM procedure. All Tukey Cramer adjusted pair-wise comparisons between Concern Levels < 0.0001.

Table 5

SNAP-IV Score Ranges and Posttest Probabilities at Various Base Rates

SNAP-IV score range	Number present	Number absent	LR	Post-test probabilities at hypothetical base rates of:		
				0.05	0.20	0.50
Parent Inattention		Gold Standard = Parent Concern				
0 – 0.6	275	759	0.4	2%	9%	29%
>0.6 – 1.2	198	81	2.8	13%	41%	74%
>1.2	284	15	21.4	53%	84%	96%
Parent Hyper/Imp						
0 – 0.6	287	719	0.5	3%	11%	33%
>0.6 – 1.2	203	113	2	10%	33%	67%
>1.2	267	23	13.1	41%	77%	93%
Teacher Inattention						
0 – 0.6	230	518	0.5	3%	11%	33%
>0.6 – 1.2	103	61	1.9	9%	32%	66%
>1.2 – 1.8	93	37	2.9	13%	42%	74%
>1.8 – 2.4	74	16	5.3	22%	57%	84%
>2.4	61	12	5.8	23%	59%	85%
Teacher Hyper/Imp						
0 – 0.6	332	554	0.7	4%	15%	41%
>0.6 – 1.2	87	48	2.1	10%	34%	68%
>1.2 – 1.8	70	18	4.5	19%	53%	82%
>1.8	68	20	3.9	17%	49%	80%
Parent Inattention		Gold standard = ADHD Diagnosis Inattentive				
0 – 0.6	7	31	0.2	1%	5%	17%
>0.6 – 1.2	27	39	0.6	3%	14%	39%
>1.2–1.8	37	28	1.2	6%	23%	55%
>1.8–2.4	23	21	1.0	5%	20%	50%
>2.4	43	8	5.0	21%	55%	83%

SNAP-IV score range	Number present	Number absent	LR	Post-test probabilities at hypothetical base rates of:			
				0.05	0.20	0.50	
Teacher Inattention							
0–0.6	25	34	0.7	3%	14%	40%	
>0.6–1.2	18	19	0.9	4%	18%	46%	
>1.2–1.8	20	15	1.2	6%	23%	55%	
>1.8–2.4	23	12	1.7	8%	30%	63%	
>2.4	14	10	1.3	6%	24%	56%	

Gold standard = ADHD Diagnosis Hyperactive/Impulsive

	Number present	Number absent	LR	0.05	0.20	0.50	
Parent Hyper/Imp							
0–0.6	8	48	0.2	1%	5%	19%	
>0.6–1.2	21	44	0.7	4%	14%	40%	
>1.2–1.8	30	38	1.1	5%	21%	52%	
>1.8	52	23	3.1	14%	44%	76%	
Teacher Hyper/Imp							
0–0.6	32	60	0.8	4%	17%	44%	
>0.6–1.2	16	19	1.3	6%	24%	56%	
>1.2–1.8	15	18	1.2	6%	24%	55%	
>1.8	13	16	1.2	6%	23%	55%	

Note. LR = Likelihood Ratio calculated as sensitivity/1 – specificity.